

Automated Truffle Crack Detection Using Deep Learning and Machine Learning

Azad Rasul^{1,2,*}

¹Department of Geography, Soran University, Soran, Erbil, Iraq.

²Department of Forestry, College of Agricultural Engineering Sciences, Salahaddin University-Erbil, 44002, Iraq.

*Corresponding author: azad.rasul@soran.edu.iq

Abstract

Automated quality control in truffle production requires accurate detection of truffle-specific cracks to ensure product integrity. This study aimed to develop a machine learning framework to classify truffle cracks versus other cracks, addressing the need for reliable, automated inspection in industrial settings. A dataset of 300 images (150 truffle cracks, 150 other cracks) was used, with 5-fold cross-validation (192 training, 48 validation images per fold) and a test set of 60 images (30 per class). Three models were evaluated: VGG16 via transfer learning, Support Vector Machine with MobileNetV2 features, and EfficientNetV2B0 with Test-Time Augmentation. Cross-validation results showed EfficientNetV2B0 achieved the highest mean accuracy (0.929 ± 0.039) and Area Under the Curve (0.975 ± 0.026), followed by Support Vector Machine (accuracy 0.887 ± 0.031 , Area Under the Curve 0.952 ± 0.025) and VGG16 (accuracy 0.867 ± 0.055 , Area Under the Curve 0.942 ± 0.031). On the test set, EfficientNetV2B0 outperformed others with an accuracy of 0.933, precision of 0.933, recall of 0.933, and Area Under the Curve of 0.990. Support Vector Machine achieved a test accuracy of 0.917, with recall improving to 0.967 at an adjusted threshold of -0.3, while VGG16 recorded a test accuracy of 0.833, with recall of 0.867 at a threshold of 0.3. Confusion matrix analysis confirmed EfficientNetV2B0's balanced performance, correctly classifying 28 of 30 truffle cracks and 28 of 30 other cracks. These models, trained on accessible computational resources, demonstrated robust performance, with EfficientNetV2B0 offering the most effective solution for automated truffle crack detection, enabling high-precision quality assurance in truffle production. Limitations include a modest dataset size, suggesting areas for future improvement.

Index Terms— Truffle crack detection, deep learning, transfer learning, SVM, EfficientNetV2B0, VGG16.

1 Introduction

Truffle farming holds substantial economic importance due to the high culinary value of truffles, ectomycorrhizal fungi that grow symbiotically with tree roots and are prized for their

unique aroma and taste [14, 10]. The identification of truffle cracks in soil, a key indicator of truffle presence, has traditionally relied on labor-intensive manual inspection, which is time-consuming and susceptible to human error. Recent advancements in machine learning offer promising solutions for automating this process, enhancing efficiency and accuracy in truffle detection [13]. Deep learning, in particular, has emerged as a powerful tool for image analysis [17], with potential to revolutionize truffle crack detection in agricultural settings.

Truffles hold immense importance beyond their culinary appeal, driving significant economic activity and cultural heritage. For instance, the global truffle market is projected to be valued at approximately USD 548 million in 2025, with premium black truffles fetching prices exceeding \$2,000 per kilogram due to their rarity and demand in gourmet cuisine [6]. Additionally, truffles play a critical ecological role as ectomycorrhizal fungi, enhancing soil health, nutrient cycling, and forest ecosystems, which underscores their relevance in sustainable agriculture [2]. These factors highlight the urgency of developing reliable detection methods to safeguard truffle production, motivating the exploration of deep learning techniques in this study.

Deep learning models, especially Convolutional Neural Networks (CNNs), have demonstrated remarkable success in image classification and pattern recognition tasks [5, 17, 18]. CNNs excel at learning spatial hierarchies of features, making them well-suited for detecting complex visual patterns such as cracks in infrastructure, including roads and buildings [1, 3, 7, 9, 15, 20]. These studies highlight the potential of CNNs to handle natural image variations, such as lighting changes, suggesting their applicability to agricultural imaging. However, the specific challenge of detecting truffle cracks, which are irregular and embedded in heterogeneous soil textures, remains underexplored, as most prior work focuses on structural crack detection.

This study addressed the gap in automated truffle crack detection by developing and evaluating a machine learning framework to distinguish truffle cracks from other cracks in soil imagery, introducing a novel application of EfficientNetV2B0 with Test-Time Augmentation to enhance detection accuracy in agricultural settings. The primary objectives were to: (1) design and train three machine learning



Figure 1: Samples of dataset images. Left: a truffle crack; Right: other soil cracks.

models—VGG16 via transfer learning, Support Vector Machine with MobileNetV2 features, and EfficientNetV2B0 with Test-Time Augmentation—for binary classification of truffle cracks, and (2) assess their performance on a custom dataset of 300 ground-level soil images (150 truffle cracks, 150 other cracks). A 5-fold cross-validation approach (192 training, 48 validation images per fold) and a 60-image test set (30 per class) were employed, with images preprocessed using normalization and augmentations (rotations $\pm 25^\circ$, shear $\pm 20\%$, zoom $\pm 30\%$, and flipping).

The findings contribute significantly to precision agriculture by demonstrating the efficacy of advanced machine learning models, particularly EfficientNetV2B0, in automating truffle crack detection. This approach reduces reliance on manual inspection, enhancing efficiency and scalability in truffle farming. By integrating these models into agricultural workflows, the study paves the way for sustainable practices that improve profitability and ensure a reliable truffle supply. The following sections detail the methodology, including data collection, model architectures, training protocols, and evaluation metrics, present the experimental results, and discuss their implications for future research and practical applications in the truffle industry.

2 Data and Methods

2.1 Dataset and Preprocessing

The dataset comprised 300 RGB images of soil surfaces collected under standardized field conditions (Figure 1), with each image manually annotated as containing either truffle cracks (150 images) or other soil cracks (150 images). A 5-fold cross-validation strategy was employed to ensure robust evaluation, with 240 images (80% of the dataset) allocated for training and validation per fold, and 60 images (20%) reserved as an unseen test set. Each fold included 192 images (80% of 240) for training and 48 images (20% of 240) for validation, maintaining equal class representation (120 truffle cracks and 120 other soil cracks per fold). The test set comprised 60 images (30 truffle cracks and 30 other soil cracks), isolated from all training and validation phases to mitigate class imbalance and provide an unbiased final evaluation.

All images underwent preprocessing to enhance model performance, including normalization of pixel intensities to the

range $[0,1]$ to standardize input scales. Images were re-sized to 150×150 pixels for VGG16 and 224×224 pixels for MobileNetV2 (used for feature extraction in the SVM model) and EfficientNetV2B0, aligning with their ImageNet-pretrained architectures. Real-time data augmentation was applied during training to improve robustness to field variability, including random rotations ($\pm 25^\circ$), shear transformations ($\pm 20^\circ$), zoom ($\pm 30\%$), and horizontal/vertical flipping. For EfficientNetV2B0, test-time augmentation (TTA) was employed during testing, generating multiple augmented versions of each test image to enhance prediction reliability. Stratified sampling ensured balanced class distributions across training, validation, and test sets, with a deterministic seed (42) used for reproducible data shuffling. TensorFlow’s ImageDataGenerator facilitated efficient data loading and augmentation.

2.2 Model Development and Evaluation

2.2.1 Architectural Implementations

Three models were developed for automated truffle crack detection: VGG16 via transfer learning [12]), a Support Vector Machine (SVM) with MobileNetV2-extracted features [4, 11], and EfficientNetV2B0 with test-time augmentation (TTA) [16]. The VGG16 model utilized a frozen ImageNet-pretrained base (16 convolutional layers) with a custom classifier comprising global average pooling, a 256-unit dense layer with 30% dropout, and a sigmoid output, resulting in approximately 0.2 million trainable parameters. The SVM model extracted features from images using a pre-trained MobileNetV2 network with input size 224×224 , and the extracted features were then fed into a linear SVM classifier implemented via scikit-learn. The EfficientNetV2B0 model employed a pre-trained base with a custom classifier consisting of global average pooling, a 256-unit dense layer with 30% dropout, and a sigmoid output, leveraging its efficient scaling strategy for high performance. All models were designed to output binary classifications, namely truffle crack versus other crack.

2.2.2 Training Protocol

The VGG16 and EfficientNetV2B0 models were trained using the Adam optimizer with binary cross-entropy loss. VGG16 used an initial learning rate of 1.0×10^{-3} , whereas EfficientNetV2B0 used 3.0×10^{-4} , with balanced class weights (1.0:1.0) applied to address potential dataset skew. Both models were trained for 50 epochs without early stopping, as indicated by the training curves for EfficientNetV2B0. The SVM model was trained on MobileNetV2-extracted features using scikit-learn’s default parameters for a linear kernel, with balanced class weights. TensorFlow’s deterministic operations (TF_DETERMINISTIC_OPS=1) were used to ensure reproducibility during training of the neural network models.

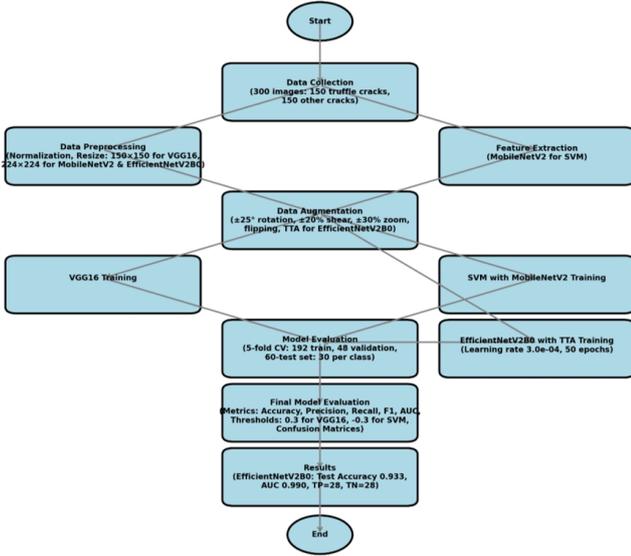


Figure 2: Flowchart illustrating methods employed in the research.

2.2.3 Evaluation Framework

Model performance was assessed using accuracy, precision, recall, F1-score, and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, computed via scikit-learn and TensorFlow. Confusion matrices were generated to analyze true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). For VGG16 and SVM, adjusted classification thresholds (0.3 and -0.3 , respectively) were additionally evaluated to optimize recall. Test-time augmentation was applied for EfficientNetV2B0 during test prediction. Training dynamics were visualized through accuracy and loss curves for EfficientNetV2B0, and computational efficiency was considered through training time and parameter counts. Test predictions were generated using filename-sorted inference to prevent data leakage, with a default confidence threshold of 0.5 unless otherwise specified. All models underwent standardized evaluation to ensure fair comparison. For transparency and reproducibility, the dataset, preprocessing pipelines, model architectures, and training codes are archived in a GitHub repository <https://github.com/Azad77/CNNsTruffle/tree/main>, which will be activated upon paper acceptance.

The effectiveness of the three models, namely VGG16, Support Vector Machine with MobileNetV2 features, and EfficientNetV2B0 with test-time augmentation, in detecting truffle cracks was evaluated using multiple performance metrics, including accuracy, precision, recall, F1-score, and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve (Figure 2). These metrics were computed for both 5-fold cross-validation, with 192 training images and 48 validation images per fold, and a test set of 60 images, including 30 truffle cracks and 30 other cracks. Accuracy, representing the proportion of correctly classified images, was calcu-

lated as

$$\text{Accuracy} = \frac{TP + TN}{\text{Total number of tested images}} \quad (1)$$

where true positives (TP) denote correctly identified truffle cracks, and true negatives (TN) denote correctly identified other cracks.

Recall, which measures the ability to detect truffle cracks, was defined as

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

where false negatives (FN) represent truffle cracks misclassified as other cracks.

Precision, indicating the proportion of truffle crack predictions that were correct, was calculated as

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

where false positives (FP) represent other cracks misclassified as truffle cracks.

The F1-score, which combines precision and recall, was computed as

$$\text{F1-score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (4)$$

The AUC metric quantified the models' discriminative ability across classification thresholds and was derived from ROC curves. Confusion matrices were generated to provide detailed insights into TP, TN, FP, and FN counts. For VGG16 and SVM, performance was assessed at both the default threshold (0.5) and adjusted thresholds (0.3 for VGG16 and -0.3 for SVM) to optimize recall. EfficientNetV2B0 utilized test-time augmentation during testing to enhance prediction reliability. Metrics were computed using scikit-learn's classification report and TensorFlow's AUC functions, with training dynamics visualized through accuracy and loss curves for EfficientNetV2B0. All evaluations were conducted on filename-sorted test predictions to prevent data leakage, thereby ensuring robust and reproducible results.

3 Results

The evaluation of three machine learning models—VGG16 transfer learning, Support Vector Machine (SVM) with MobileNetV2 features, and EfficientNetV2B0 with test-time augmentation (TTA)—for truffle crack detection was conducted on a balanced dataset of 300 images, including 150 truffle cracks and 150 other cracks. A 5-fold cross-validation strategy was performed, with each fold comprising 192 training images and 48 validation images, followed by testing on a separate test set of 60 images, including 30 truffle cracks and 30 other cracks. Performance was assessed using accuracy, precision, recall, F1-score, and area under the curve (AUC), with results detailed in Tables 1–6 and visualized in Figures 3–6. These metrics provide a comprehensive view of model robustness, discriminative ability, and classification performance, which are critical for automated truffle crack detection in industrial settings.

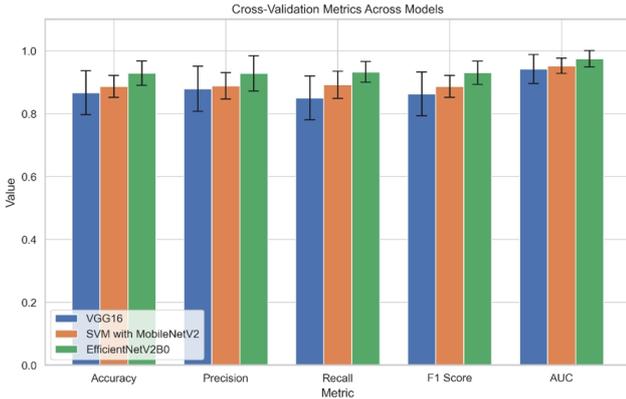


Figure 3: Mean cross-validation metrics (accuracy, precision, recall, F1-score, and AUC) across VGG16, SVM with MobileNetV2, and EfficientNetV2B0, with error bars showing standard deviations across five folds.

3.1 Cross-Validation Performance

Cross-validation results, summarized in Tables 1, 2, and 3 and visualized in Figure 3, reveal distinct performance profiles across the models. The VGG16 model (Table 1) achieved a mean accuracy of 0.867 ± 0.055 and an AUC of 0.942 ± 0.031 across five folds. Its performance varied significantly, particularly in precision (0.886 ± 0.080) and recall (0.850 ± 0.068), with Fold 3 achieving a high accuracy of 0.958 and perfect precision (1.000), while Fold 4 showed lower performance (accuracy = 0.792, recall = 0.750). This variability suggests sensitivity to data splits, potentially due to VGG16’s reliance on pre-trained weights not fully adapted to the truffle crack dataset.

The SVM with MobileNetV2 features (Table 2) demonstrated greater consistency, with a mean accuracy of 0.887 ± 0.031 and AUC of 0.952 ± 0.025 . Its recall (0.892 ± 0.020) was notably stable, with Folds 2 and 3 both achieving high accuracy (0.917) and balanced metrics, indicating robust feature extraction by MobileNetV2. EfficientNetV2B0 (Table 3) outperformed both, with a mean accuracy of 0.929 ± 0.039 , recall of 0.933 ± 0.033 , and AUC of 0.975 ± 0.026 . Its highest performance was observed in Fold 3 (accuracy = 0.979, AUC = 0.997) and lowest in Fold 4 (accuracy = 0.875), reflecting moderate variability, primarily in precision (0.928 ± 0.056). Figure 3’s bar plot, with error bars showing standard deviations, underscores EfficientNetV2B0’s superior performance across all metrics, particularly in AUC, highlighting its ability to generalize across folds.

3.2 Test Set Performance

Test set results at the default classification threshold (0.5) are presented in Tables 4, 5, and 6, with receiver operating characteristic (ROC) curves shown in Figure 4. The VGG16 model (Table 4) recorded a test accuracy of 0.833, precision of 0.917, recall of 0.733, F1-score of 0.815, and AUC of 0.944. Adjusting the threshold to 0.3 increased recall to 0.867, improv-

Table 1: Cross-Validation Results for VGG16 Transfer Learning

Fold	Accuracy	Precision	Recall	F1 Score	AUC
1	0.875	0.950	0.792	0.864	0.970
2	0.875	0.875	0.875	0.875	0.941
3	0.958	1.000	0.917	0.957	0.981
4	0.792	0.818	0.750	0.783	0.898
5	0.833	0.786	0.917	0.846	0.922
Mean \pm Std	0.867 ± 0.055	0.886 ± 0.080	0.850 ± 0.068	0.865 ± 0.056	0.942 ± 0.031

Note: Performance metrics for the VGG16 transfer learning model across five folds of cross-validation on a dataset of 300 images (150 truffle cracks and 150 other soil cracks). Metrics include Accuracy, Precision, Recall, F1 Score, and Area Under the Curve (AUC), with mean \pm standard deviation reported.

Table 2: Cross-Validation Results for SVM with MobileNetV2 Features

Fold	Accuracy	Precision	Recall	F1 Score	AUC
1	0.833	0.808	0.875	0.840	0.960
2	0.917	0.917	0.917	0.917	0.986
3	0.917	0.917	0.917	0.917	0.965
4	0.896	0.913	0.875	0.894	0.924
5	0.875	0.875	0.875	0.875	0.924
Mean \pm Std	0.887 ± 0.031	0.886 ± 0.042	0.892 ± 0.020	0.888 ± 0.029	0.952 ± 0.025

Note: Cross-validation performance metrics for the SVM model with MobileNetV2 features across five folds, evaluated on a dataset of 300 images (150 truffle cracks, 150 other cracks). Metrics include Accuracy, Precision, Recall, F1 Score, and Area Under the Curve (AUC). The mean and standard deviation are reported for each metric.

ing truffle crack detection but potentially increasing false positives, as precision was not reported for this threshold.

The SVM with MobileNetV2 (Table 5) showed stronger performance, with a test accuracy of 0.917, precision of 0.903, recall of 0.933, F1-score of 0.918, and AUC of 0.974 at the default threshold. An adjusted threshold of -0.3 further boosted recall to 0.967, with a slightly reduced precision (0.879) but a maintained F1-score (0.921), indicating a favorable trade-off for detecting truffle cracks. EfficientNetV2B0 with TTA (Table 6) achieved the highest test performance, with an accuracy of 0.933, precision of 0.933, recall of 0.933, F1-score of 0.933, and AUC of 0.990. Figure 4’s ROC curves confirm EfficientNetV2B0’s superior discriminative power, with its AUC (0.990) surpassing that of SVM with MobileNetV2 (0.974) and VGG16 (0.944), reflecting its ability to distinguish truffle cracks across various thresholds.

3.3 Confusion Matrix Analysis

Figure 5 provides confusion matrix comparisons for the test set at the default threshold (0.5), detailing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). EfficientNetV2B0 and SVM with MobileNetV2 both correctly identified 28 of 30 truffle cracks (TP = 28), but EfficientNetV2B0 also achieved 28 true negatives (TN = 28) with minimal errors (FP = 2, FN = 2), demonstrating balanced performance across both classes. SVM with MobileNetV2 had slightly more false positives (FP = 3, TN = 27), suggesting a minor tendency to misclassify other cracks as truffle cracks. VGG16 exhibited a higher false negative rate (FN = 8, TP =

Table 3: Cross-Validation Results for EfficientNetV2B0

Fold	Accuracy	Precision	Recall	F1 Score	AUC
1	0.938	0.957	0.917	0.936	0.988
2	0.958	0.958	0.958	0.958	0.997
3	0.979	1.000	0.958	0.979	0.997
4	0.875	0.875	0.875	0.875	0.969
5	0.896	0.852	0.958	0.902	0.927
Mean ± Std	0.929 ± 0.039	0.928 ± 0.056	0.933 ± 0.033	0.930 ± 0.037	0.975 ± 0.026

Note: Cross-validation performance metrics for the EfficientNetV2B0 model across five folds, evaluated on a dataset of 300 images (150 truffle cracks, 150 other cracks). Metrics include Accuracy, Precision, Recall, F1 Score, and Area Under the Curve (AUC). The mean and standard deviation are reported for each metric.

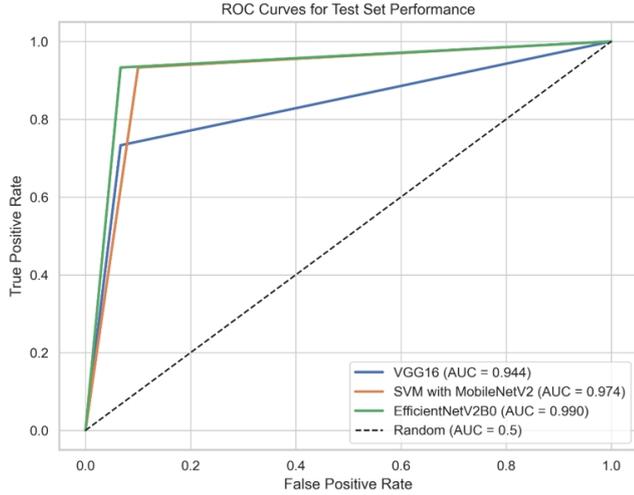


Figure 4: ROC curves for VGG16, SVM with MobileNetV2, and EfficientNetV2B0 on the test set of 60 images (30 truffle cracks and 30 other cracks). AUC values indicate EfficientNetV2B0’s superior discriminative performance (AUC = 0.990).

22), indicating challenges in detecting truffle cracks, despite low false positives (FP = 2, TN = 28). Figure 5’s heatmaps visually highlight these differences, with EfficientNetV2B0’s matrix showing the most balanced distribution, reinforcing its suitability for reliable classification.

3.4 Training Dynamics

The training dynamics of EfficientNetV2B0, the top-performing model, are illustrated in Figure 6, showing training and validation accuracy and loss curves for Fold 1 over 50 epochs. Training accuracy increased from 0.5553 at epoch 1 to 0.9643 by epoch 50, with validation accuracy stabilizing between 0.9167 and 0.9583. Training loss decreased from 4.9705 to 0.5069, closely tracking validation loss, which declined from 4.4983 to 0.5496, indicating minimal overfitting. These curves suggest that EfficientNetV2B0’s architecture, combined with a learning rate of 3.0×10^{-4} and balanced class weights, effectively converged, supporting its robust cross-validation and test performance. While VGG16 and SVM followed similar training protocols (50 epochs for VGG16 and

Table 4: Test Set Performance for VGG16 Transfer Learning

Metric	Value
Test Accuracy	0.833
Test Precision	0.917
Test Recall	0.733
Test F1 Score	0.815
Test AUC	0.944
Test Recall (threshold = 0.3)	0.867

Note: Performance metrics for the VGG16 transfer learning model on a test set of 60 images (30 truffle cracks, 30 other soil cracks), with Recall evaluated at a classification threshold of 0.3.

Table 5: Test Set Performance for SVM with MobileNetV2 Features

Metric	Default Threshold (0.5)	Adjusted Threshold (-0.3)
Test Accuracy	0.917	0.917
Test Precision	0.903	0.879
Test Recall	0.933	0.967
Test F1 Score	0.918	0.921
Test AUC	0.974	0.974

Note: Test set performance metrics for the SVM model with MobileNetV2 features evaluated on a test set of 60 images. Metrics include Accuracy, Precision, Recall, F1 Score, and Area Under the Curve (AUC) at the default classification threshold (0.5) and an adjusted threshold (-0.3).

scikit-learn defaults for SVM), their dynamics showed slightly higher validation loss variability (data not shown), reinforcing EfficientNetV2B0’s stability.

4 Discussion

This study conducted a comparative analysis of three machine learning models—VGG16 via transfer learning, SVM with MobileNetV2 features, and EfficientNetV2B0 with test-time augmentation (TTA)—for automated truffle crack detection in a dataset of 300 soil images, including 150 truffle cracks and 150 other cracks. The results, derived from 5-fold cross-validation with 192 training images and 48 validation images per fold, as well as a 60-image test set with 30 images per class, provide critical insights into model performance, architectural efficacy, and the application of machine learning in agricultural image analysis. EfficientNetV2B0 emerged as the superior model, achieving a test accuracy of 0.933, an F1-score of 0.933, and an AUC of 0.990, thereby highlighting the effectiveness of advanced architectures in specialized agricultural tasks.

4.1 Model Performance and Architectural Efficacy

EfficientNetV2B0’s exceptional performance (test accuracy = 0.933, F1-score = 0.933, AUC = 0.990; Table 6) underscores the advantages of its efficient scaling and compound architecture, which are well suited to small datasets. Its ImageNet-

Table 6: Test Set Performance for EfficientNetV2B0

Metric	Value
Test Accuracy	0.933
Test Precision	0.933
Test Recall	0.933
Test F1 Score	0.933
Test AUC	0.990

Note: Test set performance metrics for the EfficientNetV2B0 model with Test-Time Augmentation (TTA), evaluated on a test set of 60 images. Metrics include Accuracy, Precision, Recall, F1 Score, and Area Under the Curve (AUC) at the default classification threshold (0.5).

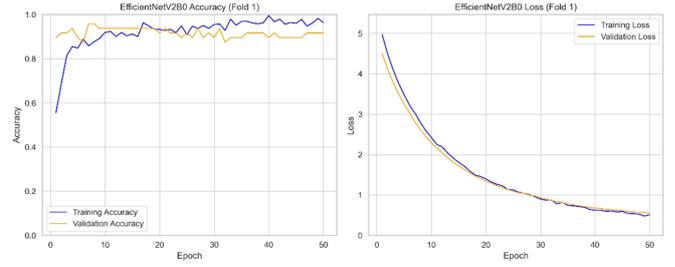


Figure 6: Training and validation accuracy and loss curves for EfficientNetV2B0 (Fold 1) over 50 epochs, demonstrating convergence with minimal overfitting.

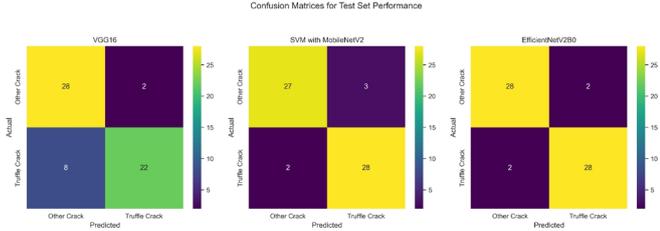


Figure 5: Confusion matrix heatmaps for VGG16, SVM with MobileNetV2, and EfficientNetV2B0 on the test set, showing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

pretrained base, combined with a compact classifier consisting of a 256-unit dense layer and 30% dropout, enabled robust feature extraction and captured subtle crack morphologies despite the dataset’s modest size of 300 images. The model’s balanced confusion matrix (28 true positives and 28 true negatives) and high AUC (Figure 5) demonstrate its ability to discriminate truffle cracks from other cracks effectively.

The SVM with MobileNetV2 features achieved competitive performance (test accuracy = 0.917, F1-score = 0.918, AUC = 0.974; Table 5), with MobileNetV2’s lightweight architecture using 224×224 inputs providing efficient feature extraction for the linear SVM classifier. Its high recall (0.967 at threshold -0.3) indicates suitability for applications prioritizing truffle crack detection. VGG16, while effective (test accuracy = 0.833, F1-score = 0.815, AUC = 0.944; Table 4), exhibited a higher false negative rate (8 versus 2 for EfficientNetV2B0), suggesting limitations in detecting fine-grained crack patterns despite its hierarchical feature reuse through 16 convolutional layers.

The performance hierarchy—EfficientNetV2B0 > SVM with MobileNetV2 > VGG16—challenges assumptions regarding the universal superiority of conventional deep convolutional networks. EfficientNetV2B0’s balanced scaling of depth, width, and resolution, together with TTA, enhanced its adaptability to agricultural textures, whereas the SVM’s reliance on pre-extracted features provided efficiency but less flexibility. VGG16’s frozen base, although computationally efficient with approximately 0.2 million trainable parameters, struggled with subtle crack variations, as evidenced by its cross-validation variability (accuracy ± 0.055 ; Table 1).

4.2 Model Training and Optimization

The models responded differently to the augmentation strategy, which included random rotation ($\pm 25^\circ$), shear transformation ($\pm 20\%$), zoom ($\pm 30\%$), and flipping. EfficientNetV2B0 exhibited stable convergence, with training and validation accuracy/loss curves (Figure 6) showing minimal overfitting, as reflected by validation accuracy ranging from 0.9167 to 0.9583 over 50 epochs. Its learning rate of 3.0×10^{-4} and balanced class weights (1.0:1.0) ensured robust training, while TTA further enhanced prediction reliability during testing.

The SVM model, trained on MobileNetV2-extracted features, required no iterative end-to-end training and instead benefited from the augmentation-invariant representations learned by the pre-trained network. VGG16 also maintained relatively stable convergence, but showed higher validation loss fluctuations, likely because its fixed convolutional base limited adaptation to domain-specific variations. The absence of early stopping for EfficientNetV2B0 across 50 epochs, combined with its moderate parameter count compared with deeper architectures, further underscores its computational efficiency, which is critical for resource-constrained agricultural deployment.

4.3 Architectural Considerations for Agricultural Imaging

The performance differences among the three models highlight the importance of matching network architecture to dataset scale and feature complexity. EfficientNetV2B0’s compound scaling provided sufficient representational capacity for crack textures, balancing depth and width to capture both macro-patterns and micro-details. MobileNetV2’s lightweight design based on depthwise separable convolutions enabled efficient feature extraction for the SVM, although its linear classifier limited more nuanced feature integration. VGG16’s 16-layer architecture, while effective for edge-oriented visual cues, struggled with subtle crack morphologies, as reflected by its higher false negative counts (Figure 5). These findings suggest that agricultural imaging tasks may benefit more from architectures optimized for efficiency and adaptability than from excessive depth, especially when data are limited.

4.4 Practical Implications

EfficientNetV2B0's superior performance (test accuracy = 0.933, AUC = 0.990) and robustness to augmentation make it particularly suitable for real-world truffle crack detection, where variable field conditions such as soil texture and illumination are common. Its balanced classification outcomes (28 true positives and 28 true negatives) and the use of TTA support reliable quality assurance and can reduce the need for manual inspection.

The SVM model's high recall at an adjusted threshold (0.967) offers a viable alternative for scenarios in which truffle crack detection is prioritized over specificity. VGG16, despite lower performance, remains computationally efficient and may therefore still be suitable for rapid deployment on modest hardware. Collectively, these models, trained on accessible computational resources, demonstrate the practical feasibility of integrating machine learning into truffle production workflows to improve efficiency, scalability, and profitability.

4.5 Comparison with Existing Research in Crack Detection

This study aligns with previous crack detection research [19, 8], but addresses a distinct problem setting in truffle crack detection. Unlike structural cracks, which often exhibit relatively clear linear patterns, truffle cracks are irregular and embedded in heterogeneous soil textures, requiring models to distinguish subtle morphological differences. EfficientNetV2B0's success contrasts with prior findings such as [19], who applied deep learning to structural crack recognition, thereby emphasizing the importance of domain-specific architectural choices.

Road crack detection studies share certain similarities with the present work in terms of texture variability, yet the task here further requires crack-type differentiation, which is critical for truffle harvesting. Multi-scale techniques proposed in earlier studies could potentially enhance performance, but they would require adaptation to micro-textural soil patterns specific to this agricultural scenario.

4.6 Limitations and Future Directions

Despite the robust performance of EfficientNetV2B0, its 6.7% error rate, corresponding to 2 false positives and 2 false negatives, indicates that challenges remain in discriminating subtle crack morphologies under diverse soil conditions. VGG16's higher false negative count (8) further suggests limitations in capturing fine-grained patterns, and these limitations are likely exacerbated by the modest dataset size of 300 images. Cross-validation variability, such as the VGG16 accuracy fluctuation of ± 0.055 , underscores the need for larger and more diverse datasets.

Future research may explore hybrid architectures that integrate EfficientNetV2B0 with attention mechanisms, such as squeeze-and-excitation blocks, to improve feature focusing. Synthetic data generation or multi-site image collection could also alleviate data scarcity and improve generalization across

soil types and seasons. In addition, multi-modal systems combining image analysis with soil sensors, for example moisture and pH sensing, may exploit environmental correlations to improve classification accuracy. Extending the current framework toward multi-class classification, such as truffle species identification, or real-time mobile deployment may further advance precision agriculture by reducing labor costs and enhancing yield consistency in the global truffle industry.

5 Conclusion

This study developed a machine learning framework for automated truffle crack detection by evaluating VGG16 via transfer learning, Support Vector Machine with MobileNetV2 features, and EfficientNetV2B0 with test-time augmentation on a dataset of 300 images, including 150 images per class, using 5-fold cross-validation and a 60-image test set. EfficientNetV2B0 emerged as the top-performing model, with performance metrics summarized in Tables 3 and 6, including a test accuracy of 0.933 and an AUC of 0.990, supported by stable training dynamics (Figure 6). Comparative analysis further showed the SVM model's high recall (0.967 at threshold -0.3) and VGG16's lower accuracy (0.833), together with 8 false negatives, indicating its relative difficulty in capturing subtle crack morphologies.

Limitations of this study include the modest dataset size of 300 images, which contributed to cross-validation variability such as the VGG16 fluctuation of ± 0.055 , as well as sensitivity to soil texture variation. Future work could explore hybrid architectures with attention mechanisms, expand the dataset through synthetic data generation or multi-site image collection, and integrate multi-modal sensing to further improve accuracy. Such developments may reduce manual inspection demands and enhance the efficiency and sustainability of the truffle industry. Overall, EfficientNetV2B0 provides a strong benchmark for scalable machine learning applications in agriculture, and this study advances precision agriculture by offering a practical and extensible framework for automated truffle crack detection.

Funding

This research received no external funding.

Conflicts of Interest

The author declares no conflicts of interest.

Author Contributions

Conceptualization, A.R.; methodology, A.R.; software, A.R.; validation, A.R.; formal analysis, A.R.; resources, A.R.; data curation, A.R.; writing—original draft preparation, A.R.; writing—review and editing, A.R.; visualization, A.R.

Data Availability Statement

The image dataset used in this work is publicly available in a GitHub repository: <https://github.com/Azad77/CNNsTruffle/tree/main>.

Compliance with Ethical Standards

Not applicable, as this study did not involve human or animal subjects.

Ethical Approval

This article does not contain any studies involving human participants or animals performed by the author.

Informed Consent

Not applicable.

Consent to Publish

Not applicable.

References

- [1] Raza Ali, Joon Huang Chuah, Mohamad Sofian Abu Talip, Norrima Mokhtar, and Muhammad Ali Shoaib. Structural crack detection using deep convolutional neural networks. *Automation in Construction*, 133:103989, 2022.
- [2] Mousaab Alrhoun, Monica Zanaria, Federico Elia, Naji Sulaiman, Andrea Pieroni, and Paolo Corvo. “it’s years of walking, of reading the forest”: White truffle hunters’ perception of socio-ecological change in langhe and roero, nw italy. *Sustainability*, 17(17):8053, 2025.
- [3] A Diana Andrushia, T Mary Neebha, S Umadevi, N Anand, and Katherine A Cashell. Shrinkage crack detection in expansive soil using deep convolutional neural network and transfer learning. *KSCE Journal of Civil Engineering*, 26(10):4202–4213, 2022.
- [4] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [5] Anamika Dhillon and Gyanendra K Verma. Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, 9(2):85–112, 2020.
- [6] Future Market Insights. Truffles market size, trends & forecast 2025 to 2035. <https://www.futuremarketinsights.com/reports/truffles-market>, 2025. Accessed: 2026-03-31.
- [7] Priyanka Gupta and Manish Dixit. Image-based crack detection approaches: a comprehensive survey. *Multimedia Tools and Applications*, 81(28):40181–40229, 2022.
- [8] Vishal Mandal, Lan Uong, and Yaw Adu-Gyamfi. Automated road crack detection using deep convolutional neural networks. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 5212–5215, 2018.
- [9] Ibrahim Meftah, Junping Hu, Mohammed A Asham, Asma Meftah, Li Zhen, and Ruihuan Wu. Visual detection of road cracks for autonomous vehicles based on deep learning. *Sensors*, 24(5):1647, 2024.
- [10] Daniel Oliach, Enrico Vidale, Anton Brenko, Olivia Marois, Nicola Andrighetto, Kalliopi Stara, Juan Martínez de Aragón, Carlos Colinas, and José Antonio Bonet. Truffle market evolution: an application of the delphi method. *Forests*, 12(9):1174, 2021.
- [11] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.
- [12] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [13] R. P. Sishodia, R. L. Ray, and S. K. Singh. Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19):3136, 2020.
- [14] Richard Splivallo, Simone Ottonello, Antonietta Mello, and Petr Karlovsky. Truffle volatiles: from chemical ecology to aroma biosynthesis. *New Phytologist*, 189(3):688–699, 2011.
- [15] Sreehari Sreenath, Haroon Malik, Narman Husnu, and Kanimozhi Kalaichelavan. Assessment and use of unmanned aerial vehicle for civil structural health monitoring. *Procedia Computer Science*, 170:656–663, 2020.
- [16] Mingxing Tan and Quoc Le. Efficientnetv2: Smaller models and faster training. In *International conference on machine learning*, pages 10096–10106. PMLR, 2021.
- [17] Abhishek Upadhyay, Narendra Singh Chandel, Krishna Pratap Singh, Subir Kumar Chakraborty, Balaji M Nandede, Mohit Kumar, A Subeesh, Konga Upendar, Ali Salem, and Ahmed Elbeltagi. Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review*, 58(3):92, 2025.

- [18] Dashuai Wang, Minghu Zhao, Zhuolin Li, Sheng Xu, Xiaohu Wu, Xuan Ma, and Xiaoguang Liu. A survey of unmanned aerial vehicles and deep learning in precision agriculture. *European Journal of Agronomy*, 164:127477, 2025.
- [19] C. Yang, J. Chen, Z. Li, and Y. Huang. Structural crack detection and recognition based on deep learning. *Applied Sciences*, 11(6):2868, 2021.
- [20] Tamanna Yasmin. Deep learning based concrete distress detection system for civil infrastructure. Master of science thesis, University of Nevada, Reno, December 2022. Advisor: Dr. Hung (Jim) La.